



# wwPDB X-ray Structure Validation Summary Report i

Feb 28, 2014 – 11:19 PM GMT

PDB ID : 2C1U  
Title : CRYSTAL STRUCTURE OF THE DI-HAEM CYTOCHROME C PEROXIDASE FROM PARACOCCUS PANTOTROPHUS - OXIDISED FORM  
Authors : Echalier, A.; Fulop, V.  
Deposited on : 2005-09-21  
Resolution : 1.95 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

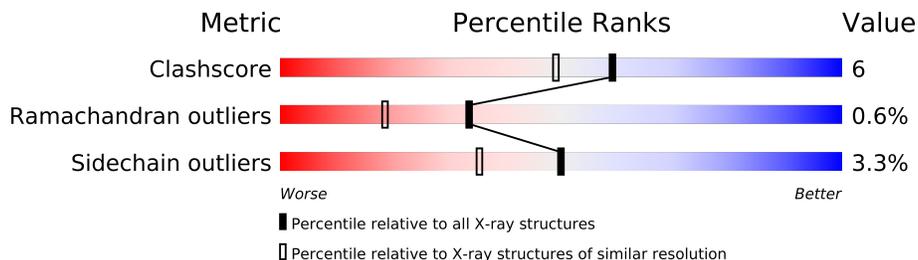
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 21963  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 1.95 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	79885	1488 (1.96-1.96)
Ramachandran outliers	78287	1475 (1.96-1.96)
Sidechain outliers	78261	1475 (1.96-1.96)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	338	
1	B	338	
1	C	338	
1	D	338	

## 2 Entry composition i

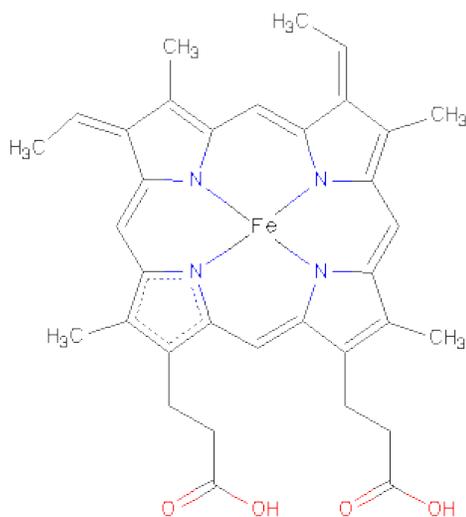
There are 4 unique types of molecules in this entry. The entry contains 11016 atoms, of which 0 are hydrogen and 0 are deuterium.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called DI-HAEM CYTOCHROME C PEROXIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	329	Total 2479	C 1566	N 415	O 486	S 12	0	0	0
1	B	324	Total 2446	C 1545	N 410	O 479	S 12	0	0	0
1	C	329	Total 2479	C 1567	N 415	O 485	S 12	0	0	0
1	D	322	Total 2431	C 1537	N 408	O 474	S 12	0	0	0

- Molecule 2 is HEME C (three-letter code: HEC) (formula:  $C_{34}H_{34}FeN_4O_4$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	Fe	N	O		
2	A	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
2	A	1	Total 43	C 34	Fe 1	N 4	O 4	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
2	B	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	B	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	C	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	C	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	D	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	D	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Ca	0	0
			1	1		
3	A	1	Total	Ca	0	0
			1	1		
3	D	1	Total	Ca	0	0
			1	1		
3	C	1	Total	Ca	0	0
			1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	208	Total	O	0	0
			208	208		
4	B	199	Total	O	0	0
			199	199		
4	C	221	Total	O	0	0
			221	221		
4	D	205	Total	O	0	0
			205	205		

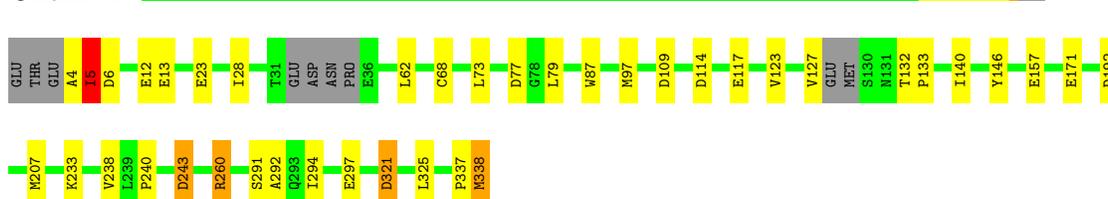
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

Note EDS was not executed.

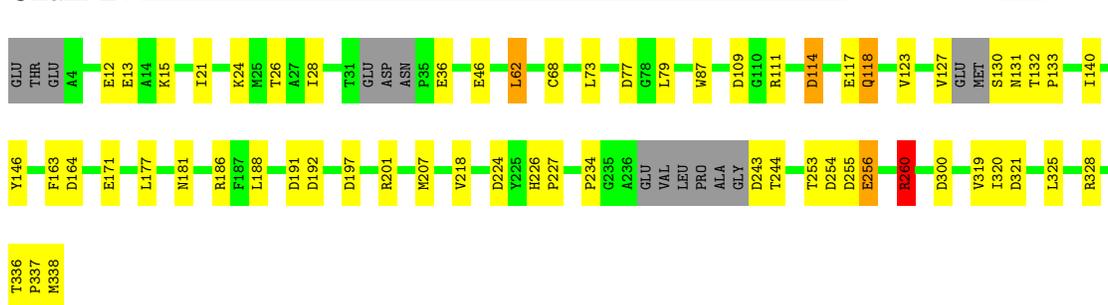
- Molecule 1: DI-HAEM CYTOCHROME C PEROXIDASE

Chain A:



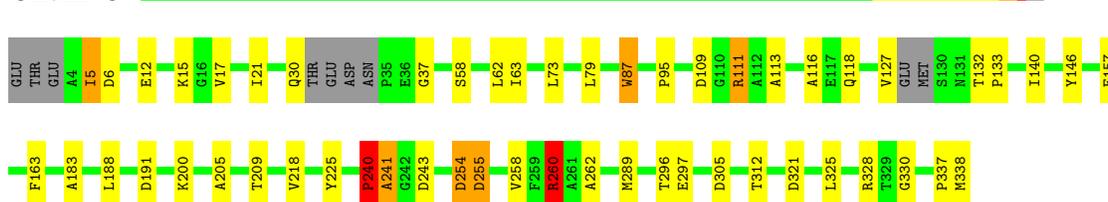
- Molecule 1: DI-HAEM CYTOCHROME C PEROXIDASE

Chain B:



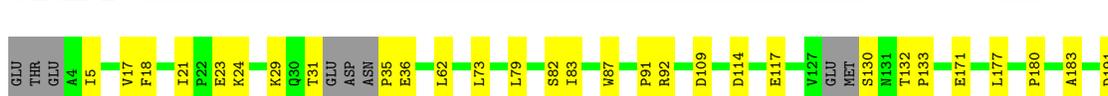
- Molecule 1: DI-HAEM CYTOCHROME C PEROXIDASE

Chain C:



- Molecule 1: DI-HAEM CYTOCHROME C PEROXIDASE

Chain D:



D192	Y218	Y225	G235	A236	VAL	LEU	PRO	ALA	GLY	ASP	THR	G245	R246	T250	D254	D255	Y258	F259	R260	E281	S291	I294	T312	E316	V319	I320	D321	L325	R328	T336	P337	M338
------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

## 4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	78.80Å 51.10Å 167.90Å 90.00° 98.20° 90.00°	Depositor
Resolution (Å)	29.88 – 1.95	Depositor
% Data completeness (in resolution range)	88.6 (29.88-1.95)	Depositor
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	REFMAC 5.1.9999	Depositor
R, $R_{free}$	0.169 , 0.220	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	11016	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

## 5 Model quality (i)

### 5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: CA, HEC

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.74	0/2535	0.89	8/3449 (0.2%)
1	B	0.72	0/2501	0.89	10/3400 (0.3%)
1	C	0.80	2/2536 (0.1%)	0.89	10/3450 (0.3%)
1	D	0.71	0/2486	0.87	10/3379 (0.3%)
All	All	0.74	2/10058 (0.0%)	0.88	38/13678 (0.3%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	12	GLU	CD-OE2	11.62	1.38	1.25
1	C	12	GLU	CD-OE1	5.39	1.31	1.25

The worst 5 of 38 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	260	ARG	NE-CZ-NH2	-10.38	115.11	120.30
1	B	328	ARG	NE-CZ-NH1	9.75	125.18	120.30
1	B	328	ARG	NE-CZ-NH2	-8.69	115.95	120.30
1	C	328	ARG	NE-CZ-NH1	7.49	124.05	120.30
1	A	321	ASP	CB-CG-OD2	7.02	124.61	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	234	PRO	Peptide

## 5.2 Close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2479	0	2392	24	0
1	B	2446	0	2358	33	0
1	C	2479	0	2393	38	0
1	D	2431	0	2347	31	0
2	A	86	0	60	5	0
2	B	86	0	60	6	0
2	C	86	0	60	7	0
2	D	86	0	60	2	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	208	0	0	2	0
4	B	199	0	0	4	0
4	C	221	0	0	3	0
4	D	205	0	0	3	0
All	All	11016	0	9730	122	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 6.

The worst 5 of 122 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:296:THR:O	1:C:297:GLU:HG2	1.48	1.12
1:D:83:ILE:CD1	1:D:91:PRO:HA	1.96	0.96
1:D:83:ILE:HD13	1:D:91:PRO:HA	1.59	0.83
1:D:83:ILE:HD11	1:D:91:PRO:CA	2.13	0.79
1:C:79:LEU:HD11	1:D:62:LEU:HG	1.69	0.74

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	323/338 (96%)	309 (96%)	12 (4%)	2 (1%)	33	19
1	B	316/338 (94%)	299 (95%)	15 (5%)	2 (1%)	33	19
1	C	323/338 (96%)	312 (97%)	8 (2%)	3 (1%)	25	11
1	D	314/338 (93%)	300 (96%)	13 (4%)	1 (0%)	50	38
All	All	1276/1352 (94%)	1220 (96%)	48 (4%)	8 (1%)	33	19

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	5	ILE
1	C	5	ILE
1	C	240	PRO
1	C	241	ALA
1	B	244	THR

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	258/267 (97%)	251 (97%)	7 (3%)	57	45
1	B	255/267 (96%)	242 (95%)	13 (5%)	33	16
1	C	258/267 (97%)	251 (97%)	7 (3%)	57	45
1	D	253/267 (95%)	246 (97%)	7 (3%)	56	43
All	All	1024/1068 (96%)	990 (97%)	34 (3%)	50	35

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	131	ASN
1	B	319	VAL
1	D	250	THR
1	B	256	GLU
1	A	338	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	30	GLN
1	B	106	GLN
1	B	131	ASN
1	C	118	GLN
1	D	181	ASN

### 5.3.3 RNA [i](#)

There are no RNA chains in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 4 are monoatomic - leaving 8 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	HEC	A	401	1	50,50,50	2.73	16 (32%)	56,82,82	2.12	17 (30%)
2	HEC	A	402	1	50,50,50	2.87	15 (30%)	56,82,82	1.82	12 (21%)
2	HEC	B	401	1	50,50,50	2.90	14 (28%)	56,82,82	2.14	14 (25%)
2	HEC	B	402	1	50,50,50	2.95	15 (30%)	56,82,82	1.86	14 (25%)
2	HEC	C	401	1	50,50,50	3.00	16 (32%)	56,82,82	2.24	18 (32%)
2	HEC	C	402	1	50,50,50	2.93	16 (32%)	56,82,82	2.07	16 (28%)
2	HEC	D	401	1	50,50,50	2.71	16 (32%)	56,82,82	2.30	18 (32%)
2	HEC	D	402	1	50,50,50	2.75	14 (28%)	56,82,82	1.79	12 (21%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	HEC	A	401	1	-	0/10/54/54	0/0/8/8
2	HEC	A	402	1	-	0/10/54/54	0/0/8/8
2	HEC	B	401	1	-	0/10/54/54	0/0/8/8
2	HEC	B	402	1	-	0/10/54/54	0/0/8/8
2	HEC	C	401	1	-	0/10/54/54	0/0/8/8
2	HEC	C	402	1	-	0/10/54/54	0/0/8/8
2	HEC	D	401	1	-	0/10/54/54	0/0/8/8
2	HEC	D	402	1	-	0/10/54/54	0/0/8/8

The worst 5 of 122 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	401	HEC	C3C-CAC	11.19	1.58	1.35
2	B	402	HEC	C3C-CAC	10.59	1.57	1.35
2	D	401	HEC	C3C-CAC	10.27	1.56	1.35
2	A	402	HEC	C3B-CAB	10.19	1.56	1.35
2	C	402	HEC	C3C-CAC	10.19	1.56	1.35

The worst 5 of 121 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	401	HEC	CBB-CAB-C3B	-8.00	105.93	128.44
2	D	401	HEC	CBB-CAB-C3B	-7.51	107.30	128.44
2	D	401	HEC	CMC-C2C-C1C	-6.93	117.96	128.62
2	A	402	HEC	CBB-CAB-C3B	-6.86	109.14	128.44

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	401	HEC	CBB-CAB-C3B	-6.56	109.96	128.44

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

EDS was not executed - this section will therefore be empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers

EDS was not executed - this section will therefore be empty.